NEW YORK UNIVERSITY
INSTITUTE OF MATHEMATICAL SCIENCES
LIBRARY
25 Waverly Place, New York 3, N. Y.

NO. 195864

IMM-NYU 249
JUNE 1958

NEW YORK UNIVERSITY
INSTITUTE OF
MATHEMATICAL SCIENCES

A Geometric Algorithm for Solving The General Linear Programming Problem

JAMES K. THURBER

NEW YORK UNIVERSITY
INSTITUTE OF MATHEMATICAL SCIENCES
LIBRARY
25 Waverly Place, New York 3, N. Y.

PREPARED UNDER
CONTRACT NO. NONR-285(32)
WITH THE OFFICE OF NAVAL RESEARCH
UNITED STATES NAVY

REPRODUCTION IN WHOLE OR IN PART
IS PERMITTED FOR ANY PURPOSE
OF THE UNITED STATES GOVERNMENT.

IMM-249
c.1

No. 195864

IMM-NYU 249
June 1958

New York University
Institute of Mathematical Sciences

A GEOMETRIC ALGORITHM FOR SOLVING THE GENERAL LINEAR PROGRAMMING PROBLEM

James K. Thurber

Prepared under the auspices of Contract Nonr-285(32) with the Office of Naval Research, United States Navy.

New York 1958

A Geometric Algorithm for Solving the General Linear Programming Problem

§1. Introduction

The development of the subject of linear programming has centered to a large extent on the problem of maximizing (or minimizing) a linear form the variables of which are subject to linear inequalities or constraints. There have been proposed many procedures for carrying out such maximizations in an effective way, foremost of which is the simplex method of G. Dantzig. All of these methods have the objectionable feature that on occasion the number of steps required, though finite, becomes excessively large. Because of this (and in the absence of any alternative) the attitude assumed by many workers in the field is that a large collection of methods should be evolved in the hope that on all occasions at least one of them will prove practicable. In this spirit, a method is proposed in this paper which is essentially a "gradient" procedure, and yet has the added feature of concluding in a finite number of steps. In some problems it works more "efficiently" than the simplex method as a consequence of the fact that instead of moving between neighboring vertices of the polyhedron of "feasible solutions" (as in the simplex method) it provides for moving across faces of this polyhedron.

§2. The linear programming problem. Let:

E_n denote euclidean n-space;

A denote an $m \times n$ matrix with rows $R_1, \ldots, R_m$;

$b = (b_1, \ldots, b_m)$ be an m-dimensional (column) vector;

$p = (p_1, \ldots, p_n)$ be an n-dimensional (row) vector.

Then the general linear programming problem may be stated as follows:

Find $x \in E_n$ such that x maximizes

$$L(x) = p \cdot x = p_1 x_1 + \ldots + p_n x_n ,$$

subject to the constraints

$$Ax \geq b .$$

(Note: The non-negativeness of the x_i's need not be among the constraints.)

§3. Further notation and its geometrical significance.

Let:

$|x|$ denote the length of a vector x E_n;

$D = \left\{ x \mid x \in E_n \text{ and } Ax \geq b \right\}$;

$D^* = \left\{ x \mid x \in D \text{ and } L(x) \geq L(u) \text{ for all } u \in D \right\}$.

The following is well known and easily verified.

Theorem 1. D and D^* are closed and convex.

If $x \in D^*$, then we call x a solution to our problem. If $x \in D$, then we call x a feasible solution.

Let $\tilde{A}$ be an rxn submatrix of A whose rows are $R_{i_1},\ldots,R_{i_r}$. Corresponding to this submatrix, let:

$\tilde{b}$ be the (column) vector $(b_{i_1},\ldots,b_{i_r})$;

$\tilde{F} = \left\{x \mid x \in D \text{ and } \tilde{A}x = \tilde{b}\right\}$;

$\tilde{V} = \left\{x \mid x \in E_n \text{ and } \tilde{A}x = 0\right\}$;

$\tilde{p}$ be the projection of p on V;

$$\tilde{q} = \begin{cases} p\cdot\tilde{p}/|\tilde{p}| & \text{if } \tilde{p} \neq 0 \\ 0 & \text{if } \tilde{p} = 0 \end{cases}.$$

(Note: We allow $r = 0$, in which case $\tilde{F} = D$, $\tilde{V} = E_n$, $\tilde{p} = p$, and $\tilde{q} = |p|$.)

Before launching our discussion of the algorithm, which will be presented in a completely algebraic context, it will be worthwhile to give geometric meaning to the symbols above, and to give a geometric description of the algorithm. For it is the geometry that motivates the method, and by interpreting our results geometrically at each stage of the discussion, the reader should have no trouble following along.

As it is well known, D is the intersection of halfspaces, and therefore a polyhedron (which may be bounded, unbounded or even empty). Then we may interpret $\tilde{F}$ as being a face of this polyhedron (which may be empty). Now suppose we have a point $x \in D$. For this x, L(x) has some value. Then what we would like to do, is to find a point $y \in D$ such that $L(y) > L(x)$. Furthermore, it would seem plausible, in moving from x to y, to do so in a manner that changes L(x) as much as possible for each unit of distance we move. Then of course, this direction

we move in will be the gradient direction p, i.e., if we move from x to a point $y = x+\lambda p$ where $\lambda > 0$, we will maximize $L(x)$ as much as is possible for each unit of distance moved. Of course, in general, we cannot move as far as we like in this direction, since we may move out of the polyhedron D. In fact we may not even be able to move at all in the gradient direction. This will occur for example if we are in certain faces of D. In this case, if x belongs to a face $\tilde{F}$, then it would be plausible to move in the gradient direction of that face. But the gradient direction of the face $\tilde{F}$ is the direction $\tilde{p}$. Furthermore the rate of change of $L(x)$ per unit distance in the direction $\tilde{p}$ is $\tilde{q}$.

We now describe our algorithm geometrically. For ease, let us assume for the moment that D is non-empty, and that $p \neq 0$. Let us further assume that we have a point $W^o \in D$. Now we find a face $\tilde{F}^o$ such that $W^o \in \tilde{F}^o$ and such that we can move in the gradient direction $\tilde{p}^o \neq 0$ from the point W^o and remain in D. We now assert:

3.1) If W^o is not already a solution, then such a face exists. If no such face exists, then $W^o \in D^*$.

Now if W^o is not a solution, we find a maximum λ such that $W^1 = W^o+\lambda\tilde{p}^o \in D$. I.e., we move as far as we can in the direction $\tilde{p}^o$. Our next assertion is:

3.2) If a maximum λ does not exist, then the problem has no solution.

This is easily seen, for in this case, $W^o+\lambda\tilde{p}^o \in D$ for all

$\lambda > 0$, and $L(W^o+\lambda\tilde{p}^o) \longrightarrow \infty$ as $\lambda \longrightarrow \infty$ (since $\tilde{p}^o \cdot p > 0$).

Now if either 3.1) or 3.2) occurs, we have solved the problem. If not, we start with the point W^1 and go through the same process. I.e., we find a face $\tilde{F}^1$ such that $W^1 \in \tilde{F}^1$ and such that we can move a positive distance in the direction $\tilde{p}^1$ ($\neq 0$). If there is no such face, then again we are at a solution. Otherwise we find a maximum μ (if it exists) such that $W^2 = W^1+\mu\tilde{p}^1 \in D$. If 3.2) applies in this case then as before no solution exists and we need not go further. Otherwise we continue the process with W^2. Finally we assert:

3.3) In a finite number of steps the process will terminate. I.e., for some integer t, we will reach a point W^t such that either

a) we cannot, for any $\tilde{F}$ containing W^t, proceed in the direction $\tilde{p}$, in which case W^t is a solution; or b) at the point W^t, 3.2) will apply, in which case there is no solution.

The rest of this paper will attempt to make the above ideas sharper, and to prove the three assertions.

Before we close this section, we will make one more point. Of the three assertions, the first one is the most subtle one. Namely, suppose we are at a point $W \in D$, such that W is not a solution. Then by the convexity of D, we know there is some direction we can move in (and still remain in D) that will improve the value of L(x). Assertion 3.1 says, that $\tilde{p}$, the gradient direction of a face $\tilde{F}$, is among these directions. Physically this is clear. For suppose p is the downward direction.

Then we can place a pebble at the point W and have it fall freely in the direction p. If we now contrain the pebble to the polyhedron D, then it is clear that it will start falling in the gradient direction of some face (which may possibly be D itself). Thus on physical grounds, assertion 3.1) should come as no suprise. However, it does remain to be proved mathematically, and our next section devotes itself to such a proof.

§4. The gradient directions as directions of motion.

Let $x,\bar{x}\in E_n$. Then we define

$$\Delta(\bar{x},x) = \begin{cases} \dfrac{L(\bar{x})-L(x)}{|\bar{x}-x|} & \text{if } \bar{x}\neq x \\ 0 & \text{if } \bar{x}=x \end{cases} .$$

We may note that if $\bar{x}\neq x$, then

$$\Delta(\bar{x},x) = p\cdot\frac{(\bar{x}-x)}{|\bar{x}-x|} .$$

We interpret $\Delta(\bar{x},x)$ to be the change of $L(x)$ per unit distance as we move from x to $\bar{x}$. We also observe that $\Delta(\bar{x},x)$ depends only on the direction $\bar{x}-x$ and not on the distance from x to $\bar{x}$. I.e., if $\bar{x} = x+\lambda y$ and $\bar{x}' = x+\lambda' y$ where $\lambda,\lambda' > 0$ and $y\neq\vec{0}$, then $\Delta(\bar{x},x) = \Delta(\bar{x}',x)$. Furthermore if $\bar{x} = x+\lambda\tilde{p}$ where $\lambda > 0$ and $\tilde{p}$ is the gradient direction of the face F, then $\Delta(\bar{x},x) = \tilde{q}$. Since p is the maximum rate of change of $L(x)$ per unit distance, the following lemma should be quite plausible.

Lemma 1. If $\bar{x},x\in F$, then $\Delta(\bar{x},x)\leq q$.

Proof. We may write p uniquely as $p = \tilde{p}+\tilde{s}$ where $\tilde{p} \in \tilde{V}$ and $\tilde{s} \in \tilde{V}^{\perp}$. Thus for any vector $y \in \tilde{V}$, $y \cdot p = y \cdot \tilde{p}$. Now if $\bar{x} = x$, then $\triangle(\bar{x},x) = 0 \leq \tilde{q}$. If $\bar{x} \neq x$. Then $\bar{x}-x/|\bar{x}-x| \in \tilde{V}$, hence

$$\triangle(\bar{x},x) = p \cdot \frac{(\bar{x}-x)}{|\bar{x}-x|} = \tilde{p} \cdot \frac{(\bar{x}-x)}{|\bar{x}-x|} \leq \tilde{p} \cdot \frac{\tilde{p}}{|\tilde{p}|} = p \cdot \frac{\tilde{p}}{|\tilde{p}|} = \tilde{q} .$$

Definition. For $x \in D$, let

$$S_x = \left\{ u \mid u \in E_n;\ |u| = 1; \text{ and for some } \lambda > 0,\ x+\lambda u \in D \right\}.$$

(Note: In the above definition λ depends on u as well as x.)

We may interpret S_x as the set of directions we may move from the point x and still remain in D. Then what we are aiming for, is to show that if $x \in D-D^*$, then there exists a gradient direction $\tilde{p}$ for some face such that $\tilde{p}/|\tilde{p}| \in S_x$.

Lemma 2. For x D, S_x is compact.

Proof. Let $I = \{1,\ldots,m\}$. Let $J = \{i \mid i \in I, \text{ and } R_i \cdot x = b_i\}$. Then for $i \in I-J$, $R_i \cdot x > b_i$. Now for $i \in I-J$, there exists $\lambda_i > 0$, such that given any unit vector $u \in E_n$, and any η such that $0 \leq \eta \leq \lambda_i$, $R_i \cdot (x+\lambda u) \geq b_i$. Let $\lambda = i\eta \min_{I-J} \lambda_i$. Now for $i \in J$, and $u \in S_x$, $R_i \cdot (x+\lambda u) \geq b_i$. Thus for this λ, $x+\lambda u \in D$ for all $u \in S_x$. Thus if $u_\ell \in S_x$ ($\ell = 1,2,3,\ldots$) has the limit point u, then $x+\lambda u_\ell \longrightarrow x+\lambda u$. But $x+\lambda u_\ell \in D$ and D is closed, hence $x+\lambda u \in D$. Therefore $u \in S_x$, showing S_x is closed. But S_x is also bounded, therefore compact.

Lemma 3. For $x \in D-D^*$, there exists $u \in S_x$ such that $u \cdot p > 0$. (Note we are assuming $D-D^*$ to be non-empty which implies $p \neq \vec{0}$.)

Proof. Since $x \in D-D^*$, there exists $y \in D$ such that $L(y) - L(x) > 0$. Let $u = (y-x)/|y-x|$. Then it is clear that u has the desired properties.

Lemma 4. For $x \in D-D^*$, there exists a unique vector $u^* \in S_x$ such that

$$u^* = \max_{u \in S_x} (p \cdot u) \quad .$$

Proof. By lemma 2, at least one such u^* exists. Now suppose there were two such vectors u_1^* and u_2^*. Then by Lemma 3, $u_1^* \neq -u_2^*$. Thus $0 < |u_1^*+u_2^*| < 2$. Now for some $\lambda_1, \lambda_2 > 0$, $x+\lambda_1 u_1^* \in D$ and $x+\lambda_2 u_2^* \in D$. Let $\lambda = \min(\lambda_1, \lambda_2)$. Then $x+\lambda u_i^* \in D$ $(i = 1,2)$. Then by the convexity of D,

$$\frac{1}{2}[x+\lambda u_1^*] + \frac{1}{2}[x+\lambda u_2^*] = x + \frac{\lambda}{2}(u_1^*+u_2^*) \in D \; .$$

Thus $(u_1^*+u_2^*)/|u_1^*+u_2^*| \in S_x$. But,

$$p \cdot \frac{u_1^*+u_2^*}{|u_1^*+u_2^*|} > p \cdot \frac{u_1^*+u_2^*}{2} = p \cdot u_1^* \; ,$$

which contradicts the definition of u_1^*.

Definition. For $x \in D-D^*$, we denote by $u^*(x)$ the unit vector whose existence and uniqueness was covered by Lemma 4.

Lemma 5. Suppose D is non-empty. Then $x \in D^*$ if and only if S_x is empty or

$$\max_{u \in S_x} (p \cdot u) \leq 0 \; .$$

Proof. Follows easily from Lemmas 2 and 3.

Theorem 2. Let $x \in D-D^*$. Then there exists a face $\tilde{F}$ (whose gradient direction is $\tilde{p}$), such that $x \in \tilde{F}$ and $u^*(x) = \tilde{p}/|\tilde{p}|$.

Proof. As in Lemma 2, let $I = \{1,\dots,m\}$, and let $J = \{i \mid R_i \cdot x = b_i\}$. Then there exists $\lambda > 0$, such that for any $i \in I-J$, and any unit vector u, $R_i \cdot (x+\eta u) > b_i$ where $0 \leq \eta \leq \lambda$. Now for some λ', $x+\lambda' u^*(x) \in D$. Let $\xi = \min(\lambda,\lambda')$ and $y = x+\xi u^*(x)$. Then $y \in D$, and for $i \in I-J$, $R_i \cdot y > b_i$. Let $J' = \{i \mid R_i \cdot y = b_i\}$. Then $J' \subset J$. Let $J' = \{i_1, i_2, \dots, i_r\}$, let A be the rxn submatrix of A whose rows are $R_{i_1},\dots,R_{i_r}$, let F be the face corresponding to $\tilde{A}$, and let $\tilde{p}$ be the gradient direction of $\tilde{F}$.

We first show that $p \neq \vec{0}$. By Lemma 3, $L(y)-L(x) = (\tilde{p}\cdot u^*(x)) > 0$. Thus by Lemma 1,

$$0 < \frac{L(y)-L(x)}{|y-x|} = \triangle(y,x) \leq \tilde{q}\ .$$

Thus by definition of $\tilde{q}$, $\tilde{p} \neq 0$.

We now show that for sufficiently small $\varepsilon > 0$, $y+\varepsilon p \in D$. Clearly we may choose ε so small that $R_i \cdot (y+\varepsilon p) > b_i$ for $i \in I-J'$. But for all positive ε, and all $i \in J'$, $R_i \cdot (y+\varepsilon\tilde{p}) = b_i$. Thus such an ε exists.

Now let $w = \xi u^*(x)+\varepsilon\tilde{p}$. Clearly $w \neq \vec{0}$, since $\tilde{p}\cdot w = \xi(p\cdot u^*(x)) + \varepsilon(p\cdot\tilde{p}) > 0$. Thus $w/|w|$ is a unit vector, and $x+|w|(\frac{w}{|w|}) \in D$. Thus $\frac{w}{|w|} \in S_x$.

Finally, we show that assumming $u^*(x) \neq \frac{\tilde{p}}{|\tilde{p}|}$, we get a contradiction. For in this case,

$$\frac{p\cdot\tilde{p}}{|\tilde{p}|} > \frac{p\cdot u^*(x)}{|u^*(x)|} = p\cdot u^*(x)$$

since $\tilde{p}$ is the gradient direction for the face $\tilde{F}$ (whether we

can move in this direction or not). Thus

$$p\cdot\frac{w}{|w|} = \frac{\xi}{|w|}(p\cdot u^{*}(x)) + \frac{\varepsilon}{|w|}(p\cdot\tilde{p}) > \frac{\xi}{|w|}(p\cdot u^{*}(x)) + \frac{\varepsilon|\tilde{p}|}{|w|}(p\cdot u^{*}(x))$$

$$= \frac{\xi+\varepsilon|\tilde{p}|}{|\xi u^{*}(x)+\varepsilon p|}(p\cdot u^{*}(x)) \geq \frac{\xi+\varepsilon|\tilde{p}|}{\xi|u^{*}(x)|+\varepsilon|p|}(p\cdot u^{*}(x)) = p\cdot u^{*}(x).$$

Thus $p\cdot(w/|w|) > p\cdot u^{*}(x)$ which contradicts our definition of $u^{*}(x)$ (since $(w/|w|)\in S_x$). Thus we must conclude that $u^{*}(x) = \tilde{p}/|\tilde{p}|$.

Corollary 1. For $x \in D-D^{*}$, there exists $y \in D$ such that $y-x = \lambda\tilde{p}$ for some $\tilde{p}$ and some $\lambda > 0$. Furthermore for this y, $\triangle(y,x) > 0$.

§5. The geometric algorithm.

We now give the algebraic formulation of the algorithm which we described geometrically in Section 3. So suppose we are given the constraints

$$Ax \geq b$$

and we wish to maximize

$$L(x) = p\cdot x\ .$$

We further make the assumption that we are given

$$w^{0} \in D\ .$$

Of course this is a strong assumption, but in Section 7, we will illustrate (using this geometric algorithm) how one obtains such a w^{0} (if it exists). We of course also know the directions $\tilde{p}$ corresponding to the faces $\tilde{F}$. (These directions,

depend only on the matrix A and the vector p. A discussion of how they are computed will be left for Section 8.)

Starting with w^o, we now define a sequence of points, either finite or infinite as follows: (We use induction)

Suppose w^k is already defined ($k \geq 0$). Let $\tilde{F}_1, \ldots, \tilde{F}_\ell$ be all the faces containing w^k. Let $\tilde{p}_1, \ldots, \tilde{p}_\ell$ be the gradient directions for these faces. Then either

I) $\tilde{p}_j/|\tilde{p}_j| \notin S_{w^k}$ for $j = 1, \ldots, \ell$ ($\tilde{p}_j \neq 0$).

or

II) There exists $\bar{j}$ ($1 \leq \bar{j} \leq \ell$) such that $\tilde{p}_{\bar{j}} \neq 0$ and $\tilde{p}_{\bar{j}}/|\tilde{p}_{\bar{j}}| \in S_{w^k}$.

If II) holds, let $L = \left\{ i \,|\, R_i \cdot \tilde{p}_{\bar{j}} < 0 \right\}$. Then:

a) If I) holds, terminate the sequence at w^k.

b) If II) holds, and L is empty, terminate the sequence at w^k.

c) If II) holds, and L is non-empty, define

$$w^{k+1} = w^k + \lambda \tilde{p}_{\bar{j}}$$

where

$$\lambda = \min_{i \in L} \left\{ \frac{b_i - R_i \cdot w^k}{R_i \cdot \tilde{p}_{\bar{j}}} \right\} .$$

Theorem 3. If the sequence terminates at w^k, because of a), then w^k is a solution.

Proof. Follows from Theorem 2 and its corollary.

Theorem 4. If the sequence terminates at w^k because of b), then there is no solution.

Proof: Since II) holds, $\tilde{p}_{\bar{j}} \neq 0$. Thus $\tilde{p}\cdot\tilde{p}_{\bar{j}} > 0$, hence $L(w^k+\mu\tilde{p}_{\bar{j}}) \longrightarrow \infty$ as $\mu \longrightarrow \infty$. But under b), $R_i\cdot\tilde{p}_{\bar{j}} \geq 0$, hence $R_i\cdot(w^k+\mu\tilde{p}_{\bar{j}}) \geq b_i$ for all i and $\mu \geq 0$. Thus $w^k+\mu\tilde{p}_{\bar{j}}) \in D$ for all $\mu \geq 0$, hence the problem has no solution.

In view of our definition of the sequence $w^o,w^1,w^2,\ldots,$ and in view of the last two theorems, it follows, that if we can give a procedure for constructing the sequence, and if we can prove that the sequence always terminates in a finite number of steps, then we will have an algorithm for solving the programming problem. We will accomplish this shortly, but first we shall prove a result which will shed some light on how this procedure is comparable with our geometrical description. It is clear that if w^k and w^{k+1} are two consecutive points in our sequence, where $w^{k+1} = w^k+\lambda\tilde{p}_{\bar{j}}$, then we did move in the gradient direction of some face to obtain w^{k+1} from w^k. We would like now to show, that we actually have moved as far as possible in the direction $\tilde{p}_{\bar{j}}$.

Theorem 5. Let w^k and $w^{k+1} = w^k+\lambda\tilde{p}_{\bar{j}}$ be two consecutive points in the sequence. Then $\lambda > 0$, and for $\mu > \lambda$, $w^k+\mu\tilde{p}_{\bar{j}} \notin D$.

Proof. Since $\tilde{p}_{\bar{j}}/|\tilde{p}_{\bar{j}}| \in S_{w^k}$, there exists $\eta > 0$, such that $w^k+\eta\tilde{p}_{\bar{j}} \in D$. Now consider some $i \in L$. Then $R_i\cdot\tilde{p}_{\bar{j}} < 0$. But for this i, $R_i\cdot(w^k+\eta\tilde{p}_{\bar{j}}) \geq b_i$. Thus $\eta\,(R_i\cdot\tilde{p}_{\bar{j}}) \geq b_i-R_i\cdot w^k$. Hence

$$\eta \leq \frac{b_i-R_i\cdot w^k}{R_i\cdot\tilde{p}_{\bar{j}}} .$$

But this holds for all $i \in L$. Thus

$$0 < \eta \le \min_{i \in L} \left\{ \frac{b_i - R_i \cdot w^k}{R_i \cdot \tilde{p}_{\bar{j}}} \right\} = \lambda \ .$$

Hence $\lambda > 0$.

Now suppose $\mu > \lambda$. Let $\bar{i} \in L$ be such that

$$\lambda = \frac{b_{\bar{i}} - R_{\bar{i}} \cdot w^k}{R_{\bar{i}} \cdot \tilde{p}_{\bar{j}}} \ .$$

Since $R_{\bar{i}} \cdot \tilde{p}_{\bar{j}} < 0$,

$$\mu(R_{\bar{i}} \cdot \tilde{p}_{\bar{j}}) < \lambda(R_{\bar{i}} \cdot \tilde{p}_{\bar{j}}) = b_{\bar{i}} - R_{\bar{i}} \cdot w^k \ .$$

Thus

$$\mu(R_{\bar{i}} \cdot \tilde{p}_{\bar{j}}) + R_{\bar{i}} \cdot w^k < b_{\bar{i}}$$

thus

$$R_{\bar{i}} \cdot (w^k + \mu p_{\bar{j}}) < b_{\bar{i}} \ .$$

Hence $w^k + \mu \tilde{p}_{\bar{j}}$ does not satisfy the $\bar{i}^{th}$ constraint.

Corollary 2. $|w^{k+1} - w^k| > 0$.

Corollary 3. If w^{k-1}, w^k, w^{k+1} are three consecutive points of the sequence, then they are non-collinear. I.e.,

$$|w^{k+1} - w^k| + |w^k - w^{k-1}| > |w^{k+1} - w^{k-1}|$$

§6. The finiteness of the algorithm.

Let us enumerate all the faces F of the polyhedron D, for which $\tilde{p} \ne 0$ (and hence $\tilde{q} > 0$). Namely let these faces be

$$\tilde{F}_1, \tilde{F}_2, \ldots, \tilde{F}_h \qquad (h \le 2^m)$$

where $0 < \tilde{q}_1 \le \tilde{q}_2 \le \ldots \le \tilde{q}_h$.

Let the corresponding gradient directions be

$$\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_h \,.$$

Now suppose $w^o \in D$. Let

$$w^o, w^1, w^2, \ldots \quad \text{(finite or infinite)}$$

be the sequence defined inductively in Section 5. Let

$$w^k = w^{k-1} + \lambda_k p_{j_k} \qquad (k = 1,2,\ldots)$$

where $1 \leq j_k \leq h$. Then

Lemma 6. If $j_k = j_\ell$ $(k < \ell)$, then there exists s such that $k < s < \ell$, and $\tilde{q}_{j_s} < \tilde{q}_{j_k}$.

Proof. Let $\tilde{F} = \tilde{F}_{j_k}$, $\tilde{p} = \tilde{p}_{j_k}$, $\tilde{q} = \tilde{q}_{j_k}$. Now $w^k \in \tilde{F}$ and $w^{\ell-1} \in \tilde{F}$. Let $L = |w^{\ell-1} - w^k|$, $L_1 = |w^{k+1} - w^k|$, $L_2 = |w^{k+2} - w^{k-1}|, \ldots,$ $L_{\ell-k-1} = |w^{\ell-1} - w^{\ell-2}|$. Then by Corollary 3,

$$L < L_1 + L_2 + \ldots + L_{\ell-k-1} \,.$$

By Lemma 1, $\triangle(w^{\ell-1}, w^k) \leq \tilde{q}$. But

$$L\triangle(w^{\ell-1}, w^k) = \tilde{q}_{j_{k+1}} L_1 + \tilde{q}_{j_{k+2}} L_2 + \ldots + \tilde{q}_{j_{\ell-1}} L_{\ell-k-1} \,.$$

Now suppose the lemma is false, i.e., $\tilde{q}_{js} \geq \tilde{q}$ for $k < s < \ell$. Then

$$\tilde{q}L \geq \triangle(w^{\ell-1}, w^k)L \geq \tilde{q}(L_1 + \ldots + L_{\ell-k-1}) > \tilde{q}L \,.$$

Thus we have a contradiction.

Now let $\tilde{p}_{j1}, \tilde{p}_{j2}, \ldots,$ be the sequence of directions (defined as before) used to obtain the sequence $w^o, w^1, \ldots$. Let $N(i)$ be the number of times $\tilde{p}_i$ occurs in the above sequence ($N(i)$ could be ∞). Then

Theorem 6. $N(i) \leq 2^{i-1}$.

Proof. Suppose that $\sum_{\nu=1}^{i-1} N(\nu) = K$ is finite. Suppose further that $\tilde{p}_i$ occured K+2 times. We will then get a contradiction. For suppose $\tilde{p}_i$ occured in the sequence in the

$$\ell_1, \ell_2, \ell_3, \ldots, \ell_{K+2}$$

places. Then some $\tilde{p}_s$ with $s < i$ would have to occur between the ℓ_t and the ℓ_{t+1} terms of the sequence by Lemma 6. Thus $\sum_{\nu=1}^{i-1} N(\nu) \geq K+1$ which is a contradiction. Hence

$$N(i) \leq \sum_{\nu=1}^{i-1} N(\nu)+1 .$$

But $N(1) \leq 1$ by Lemma 7. Thus a simple induction argument gives the result.

Corollary. There are at most 2^h terms in the sequence $w^o, w^1, \ldots$. (Thus there are at most 2^{2^n} terms in the sequence.)

§7. Obtaining a feasible solution.

Let us assume that given $w^o \in D$, we know how to compute $w^1, w^2, \ldots$ and we are able to determine when the sequence terminates. Then the whole algorithm would depend on obtaining $w^o \in D$. But to find such a vector, we essentially alter the problem as follows. Let B be an m-dimensional column vector all of whose components are 1. Let x_{n+1} be a new variable. Then consider the problem of maximizing

$$L(x_1, x_2, \ldots, x_{n+1}) = -x_{n+1} ,$$

subject to the constraints

$$Ax+Bx_{n+1} \geq b \ , \qquad\qquad x_{n+1} \geq 0 \ .$$

It is then clear that a feasible solution to the altered problem exists, namely, $x_i = 0$ $(i = 1,\ldots,n)$, $x_{n+1} = \max(0,b_1,\ldots,b_m)$. It also is clear that a solution to the new problem exists, since $L_1(x_1,\ldots,x_{n+1})$ is bounded above by zero. We may then apply our algorithm to the new problem and obtain a solution $\bar{x}_1,\ldots,\bar{x}_n,\bar{x}_{n+1}$. Then $x = (\bar{x}_1,\ldots,\bar{x}_n)$ will be feasible for our original problem, if and only if $\bar{x}_{n+1} = 0$.

§8. The computations.

To carry out the computational scheme, we must be able to

1) compute $\widetilde{p}$ for a face $\widetilde{F}$;

and

2) decide for a given $\widetilde{p} \neq 0$, whether or ont $\widetilde{p}/|\widetilde{p}| \in S_w$ for some $w \in D$.

That these can actually be carried out is shown in the following two results.

Theorem 7. $\widetilde{p} = p-\widetilde{A}^{\perp}\mu$ where μ is any solution to the r equations

$$\widetilde{A}p = \widetilde{A}\widetilde{A}^{\perp}\mu \ .$$

Furthermore this set of equations always has a solution.

Proof: We may write $p = \widetilde{p}+\widetilde{s}$ where $\widetilde{p} \in \widetilde{V}$ and $\widetilde{S} \in \widetilde{V}^{\perp}$. Furthermore this decomposition is unique. Also $\widetilde{V}$ is the set of vectors in E_n, orthogonal to all the rows of $\widetilde{A}$. Thus $\widetilde{V}^{\perp}$ is

generated by the vectors $R_{i_1},\dots,R_{i_r}$. Thus we may write $\tilde{S} = \mu_1 R_{i_1}+\dots+\mu_r R_{i_r} = A^{\perp}\mu$ where $\mu = (\mu_1,\dots,\mu_r)$. Then $\tilde{A}p = \tilde{A}(\tilde{p}+\tilde{s}) = \tilde{A}\tilde{s} = \tilde{A}\tilde{A}^{\perp}\mu$. Thus the equations have at least one solution.

Now let $\mu = (\mu_1,\dots,\mu_r)$ be any solution to the equations $\tilde{A}p = \tilde{A}\tilde{A}^{\perp}\mu$. Then let $\tilde{s} = \tilde{A}^{\perp}\mu = \mu_1 R_{i_1}+\dots+\mu_r R_{i_r}$. Then $\tilde{s}\in\tilde{V}^{\perp}$. Now $\tilde{A}(p-\tilde{s}) = \tilde{A}p-\tilde{A}\tilde{A}^{\perp}\mu = 0$. Thus $p-\tilde{s}\in\tilde{V}$. But $p = (p-\tilde{s})+\tilde{s}$, where $p-\tilde{s}\in\tilde{V}$ and $\tilde{s}\in\tilde{V}^{\perp}$, has a unique such decomposition. Thus $\tilde{p} = p-\tilde{s} = p-\tilde{A}^{\perp}\mu$.

Theorem 8. Let $w\in D$. Let $\tilde{p}$ be the gradient direction for $\tilde{F}$ (where $\tilde{p}\neq\vec{0}$). Let $L = \{i \mid R_i\cdot\tilde{p} < 0\}$ be non-empty. Let

$$\lambda = \min_{i\in L}\left\{\frac{b_i-R_i\cdot w}{R_i\cdot\tilde{p}}\right\}.$$

Then $\tilde{p}/|\tilde{p}|\in S_w$ if and only if $\lambda > 0$.

Proof. Suppose $\lambda > 0$. Then $\bar{w} = w+\lambda\tilde{p}$ is distinct from w. Now for $i\notin L$ $R_i\cdot(w+\lambda\tilde{p}) \geq R_i\cdot w \geq b_i$. Further for $i\in L$, $R_i\cdot(\lambda\tilde{p}) \geq b_i-R_i\cdot w$. Thus $R_i\cdot(w+\lambda\tilde{p}) \geq b_i$. Thus $\bar{w}\in D$. Hence $w+(\lambda|\tilde{p}|)\tilde{p}/|\tilde{p}|\in D$ where $\lambda|\tilde{p}| > 0$. Thus $\tilde{p}/|\tilde{p}|\in S_w$.

The implication in the other direction follows from Theorem 6.

Theorems 7 and 8 then give the computational procedure for finding the sequence $w^0, w^1, \dots$. In view of the fact that we now know how to compute the sequence, and in view of the fact that it terminates, we now are able to determine a solution or to decide that a solution does not exist. We now illustrate the algorithm with an example.

§9. An example.

Maximize

$$L(x,y,z) = -2x-y+2z \ ,$$

subject to the constraints,

$$\begin{cases} 2x \quad\quad -z \geq -6 & (1) \\ -x +11y -5z \geq -30 & (2) \\ -13x -22y +z \geq -60 & (3) \\ \quad y \geq 0 & (4) \end{cases}$$

In this case $\tilde{p} = (-2,-1,2)$ and it can be verified that $w^o = (3,1,2)$ is a feasible solution for which a strict inequality holds in each constraint.

Now we let

$$w^1 = w^o + \lambda_1 \tilde{p}_1 \ .$$

But in this case, $\tilde{p}_1 = \tilde{p}$ since no equality holds. Now

$\tilde{p}_1 \cdot R_1 = -6$, $\tilde{p}_1 \cdot R_2 = -19$, $\tilde{p}_1 \cdot R_3 = 50$, $\tilde{p}_1 \cdot R_4 = -1$.

Thus

$$L_1 = \{ i \mid \tilde{p}_1 \cdot R_i < 0 \} = \{ 1,2,4 \}$$

then

$$\lambda_1 = \min_{i \in L_1} \left\{ \frac{b_i - R_i \cdot w^o}{R_i \cdot \tilde{p}_1} \right\} = \frac{b_4 - R_4 \cdot w^o}{R_4 \cdot \tilde{p}_1} = 1 \ .$$

Hence $w^1 = w^o + \tilde{p}_1 = (1,0,4)$.

Now w^1 satisfies the equality in the fourth constraint. Thus we let

$$A = (0,1,0) \ .$$

Now $\tilde{p}_2 = \tilde{p}-\mu R_4$ where μ satisfies

$$\tilde{A}\tilde{p} = \tilde{A}\tilde{A}^{\perp}\mu \; .$$

Thus $-1 = \mu$, hence $\tilde{p}_2 = (-2,0,2)$. Then $w^2 = w^1+\lambda_2\tilde{p}_2$. In order to compute λ_2, we first compute

$$\tilde{p}_2 \cdot R_1 = -6 \; , \; \tilde{p}_2 \cdot R_2 = -8 \; , \; \tilde{p}_2 \cdot R_3 = 28 \; , \; \tilde{p}_2 \cdot R_4 = 0 \; .$$

Then $L_2 = \left\{ i \,|\, R_i \cdot \tilde{p}_2 < 0 \right\} = \left\{ 1,2 \right\}$, and

$$\lambda_2 = \min_{i \in L_2} \left\{ \frac{b_i - R_i \cdot w^1}{R_i \cdot \tilde{p}_2} \right\} = \frac{b_1 - R_1 \cdot w^1}{R_1 \cdot \tilde{p}_2} = 2/3 \; .$$

Thus $w^2 = (-1/3,0,16/3)$.

Now w^2 satisfies constraint (1). Therefore we try

$$\tilde{A} = (2,0,-1) \; .$$

We then set

$$\tilde{p}_3 = \tilde{p}-\mu R_1$$

where μ satisfies

$$\tilde{A}p = \tilde{A}\tilde{A}^{\perp}\mu \; .$$

Thus $\mu = -6/5$. Hence

$$\tilde{p}_3 = (\tfrac{2}{5},-1,\tfrac{4}{5}) \; .$$

Thus $w^3 = w^2+\lambda_3\tilde{p}_3$. Now

$$\tilde{p}_3 \cdot R_1 = 0 \; , \; \tilde{p}_3 \cdot R_2 = -\frac{77}{5} \; , \; \tilde{p}_3 \cdot R_3 = \frac{88}{5} \; , \; \tilde{p}_3 \cdot R_4 = -1 \; .$$

Thus $L_3 = \left\{ i \,|\, R_i \cdot \tilde{p}_3 < 0 \right\} = \left\{ 2,4 \right\}$, and

$$\lambda_3 = \min_{i \in L_3} \left\{ \frac{b_i - R_i \cdot w^2}{R_i \cdot \tilde{p}_3} \right\} = \frac{b_4 - R_4 \cdot w^2}{R_4 \cdot \tilde{p}_3} = 0 \; .$$

Thus we cannot move in the direction $\tilde{p}_3$ chosen. Hence we try

to move in a different direction. Knowing that w^2 satisfies both (1) and (4), we now try

$$\tilde{A} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} .$$

We then set

$$\tilde{p}_3 = \tilde{p} - \mu_1 R_1 - \mu_2 R_2$$

where

$$\tilde{A}p = \tilde{A}\tilde{A}^{\perp} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

or

$$\begin{pmatrix} -6 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} .$$

Thus $\mu_1 = -\frac{6}{5}$ and $\mu_2 = -1$. Then

$$\tilde{p}_3 = (\tfrac{2}{5}, 0, \tfrac{4}{5}) .$$

For this $\tilde{p}_3$,

$$R_1 \cdot \tilde{p}_3 = 0 \ , \ R_2 \cdot \tilde{p}_3 = -\frac{22}{5} \ , \ R_3 \cdot \tilde{p}_3 = -\frac{22}{5} \ , \ R_4 \cdot \tilde{p}_3 = 0 \ .$$

In this case,

$$L_3 = \{2, 3\}$$

and

$$\lambda_3 = \frac{b_2 - R_2 \cdot w^2}{R_2 \cdot \tilde{p}_3} = 5/6 .$$

Thus $w^3 = w^2 + \lambda_3 \tilde{p}_3 = (0,0,6)$.

Now w^3 satisfies the equality in (1), (2), (4). Thus there are eight choices of $\tilde{A}$. Now for most of these A we can find by trial, that we cannot move in the corresponding direction $\tilde{p}$. But if we take

$$\tilde{A} = \begin{pmatrix} 2 & 0 & -1 \\ -1 & 11 & -5 \end{pmatrix} ,$$

Then we find that $\tilde{p}_4 = (\frac{1}{6},\frac{1}{6},\frac{1}{3})$. Also

$$R_1\cdot\tilde{p}_4 = 0 \ , \ R_2\cdot\tilde{p}_4 = 0 \ , \ R_3\cdot\tilde{p}_4 = -\frac{11}{2} \ , \ R_4\cdot\tilde{p}_4 = \frac{1}{6} \ .$$

Thus $L_4 = \left\{3\right\}$ and $\lambda_4 = 12$. Hence $w^4 = w^3+\lambda_4\tilde{p}_4 = (2,2,10)$.

It is then found that w^4 satisfies the equalities in (1), (2), (3). Thus there are again eight choices of A. Corresponding to these A, we have eight direction $\tilde{p}$. If we now try to move in any of these eight directions, we find it impossible. Thus by Theorem 2, w^4 must be the solution.

§10. Motion in the "best possible" direction.

In Section 5, we showed that given $w^o \in D$, we may define a sequence (which we proved to be finite),

$$w^o, w^1, \ldots, w^t$$

where

$$w^k = w^{k-1}+\lambda_k\tilde{p}_{j_k} \qquad (k = 1,\ldots,t).$$

However the sequence is not uniquely defined, since at each stage we may have a choice in the direction p_{j_k}. Now suppose we specify that

$$\tilde{p}_{j_k}/|\tilde{p}_{j_k}| = u^*(w^{k-1}) \ .$$

That this is possible is clear by Theorem 2. If we make this choice at each stage, then the above sequence is uniquely defined. Furthermore we have the following:

Theorem 9. If $\tilde{p}_{j_k}/|\tilde{p}_{j_k}| = u^*(w^{k-1})$ for $k = 1,\ldots,t$, then

$q_{j_k} > q_{j_{k+1}}$ for $k = 1,\dots,t-1$.

Proof. Suppose this were not the case. Then there exists ℓ such that $1 \le \ell \le t-1$ and $q_{j_\ell} \le q_{j_{\ell+1}}$. Let $L_1 = |w^{\ell}-w^{\ell-1}|$, $L_2 = |w^{\ell+1}-w^{\ell}|$ and $L = |w^{\ell+1}-w^{\ell-1}|$.

Then by definition of $u^*(w^{\ell-1})$,

$$\tilde{p}\cdot u^*(w^{\ell-1}) \ge \triangle(w^{\ell+1},w^{\ell-1}) = \frac{q_{j_\ell}L_1+q_{j_{\ell+1}}L_2}{L} \ge \frac{q_{j_\ell}(L_1+L_2)}{} > q_{j_\ell}$$
$$= \tilde{p}\cdot u^*(w^{\ell-1}) .$$

But this is clearly a contradiction, hence we conclude the truth of the theorem.

Corollary 4. If $p_{j_k}/|p_{j_k}| = u^*(w^{k-1})$, then $N(i) \le 1$.

Thus if at each stage in the sequence, we move in the direction $u^*(w^{k-1})$, which is the "best possible" direction of motion (in the sense that we increase $L(x)$ as much as possible per unit distance), then Corollary 4 says that we move through each face at most once. Of course this is much better than the result of Theorem 6. However in computing $u^*(w^{k-1})$ at each stage, is in general tedious. The above situation corresponds to a falling pebble constrained to a polyhedra, and Corollary 4 states that the pebble will pass through each face at most once, which substantiates our physical intuition.

OCT 3 0 1959

Date Due

MAR 18 '68

PRINTED IN U. S. A.

NYU c. 1
IMM-
249 Thurber.
AUTHOR
A geometric algorithm for
TITLE solving the general linear
programming problem

DATE DUE	BORROWER'S NAME	ROOM NUMBER
MAR [illegible] '68	Spillers	430

N. Y. U. Institute of
Mathematical Sciences
25 Waverly Place
New York 3, N. Y.

INSTITUTE OF MATHEMATICAL SCIENCES
New York 3, N. Y.